

PROCESSING A SPEECH FRAME IN A RADIO SYSTEM

FIELD

5 [0001] The invention relates to a method of processing a speech frame in a radio system, a radio system, a mobile station of a radio system, and a network of a radio system.

BACKGROUND

10 [0002] The speech to be transmitted in a radio system is coded in a transmitter by means of a speech coder. An example of a speech coder is the ACELP (Algebraic Code Excited Linear Predictive) coder of the TETRA (Terrestrial Trunked Radio) system. The Tetra system is described in John Dunlop, Demessie Girma, James Irvine: *Digital Mobile Communications and the TETRA System*, John Wiley & Sons Limited 1999, ISBN 0-471-98792-1, incorporated as reference herein. The receiver comprises a speech decoder converting the speech information received back into speech that is
15 understood by humans. If the speech coder itself comprises a speech decoder, the device is called a speech codec.

20 [0003] The digital speech information is channel-coded in a transmitter to prevent the adverse effects caused by the noise on the radio path. Channel codes may be error-detecting and/or error-correcting codes. In principle, the channel codes can be divided into block codes and convolutional codes. Both coding methods can be used simultaneously. One of the block codes is cyclic redundancy check, CRC. Further, convolutional coding and different modifications thereof, such as punctured convolutional coding, are typically used. A typical code rate for convolutional coding, in other words the
25 ratio of the number of user data bits to the coded data bits of a channel, is for example 1/2 or 1/3. Interleaving is often used in addition to channel coding. In interleaving, the successive bits are mixed with each other for a longer period of time in such a way that a momentary fade on the radio path would not be sufficient to make the radio signal in this period unidentified but that the errors
30 in the signal could still be eliminated by decoding the channel coding.

35 [0004] Information can be not only channel-coded but also encrypted to prevent tapping. Encryption is usually implemented in such a way that an encryption mask is created by means of an encryption algorithm, using certain input parameters, such as an encryption key, the encryption algorithm being connected to the information to be transmitted by means of an XOR

operation (a logic exclusive OR operation). Encryption may be performed either prior to or after channel coding. There may be encryption on several levels, for example encryption over a radio interface, and point-to-point encryption for example between two mobile stations. The encryption can be
 5 decrypted in the receiver by using the same encryption algorithm that was used in the transmission, as far as the input parameters are also the same.

[0005] An error-detecting channel code, for example computation of a cyclic redundancy check, usually expresses the bit errors having entered the channel-coded speech frame on the radio path, whereby a bad frame
 10 indication, BFI, is sent to the speech decoder, from which indication the speech decoder detects that it is not advantageous to attempt to decode the speech frame concerned. Thus, the speech decoder generally utilizes a method in which the parameters of the preceding unbroken frame is used, the gain being at the same time attenuated.

[0006] However, in some cases, when the bits of a speech frame suitably invert on the radio path, the speech frame no longer contains acceptable data that could be decoded by a speech decoder, although on the basis of the channel code decoding, the speech frame is free of defects. Since there are a very large number of possible bit combinations in a speech frame,
 20 for example $2^{137} - 1$ different combinations in TETRA, it is in practice impossible to test them all. Thus, attempts to decode a defective speech frame may cause a very strong defective signal to exit from the output of the speech decoder, causing an unpleasant audio shock in the hearer's ears and at worst leading to a hearing defect.

[0007] The problem is made worse by the fact that in the TETRA system, for example, the intensity of sound in the loudspeaker of mobile stations is, according to clients' wishes, in the range of the maximum values allowed to maximize the audibility in noisy environments. Thus, even if the loudspeaker circuit has its own restriction of sound intensity, for instance 100
 30 or even 110 decibels, this does not solve the problem. The problem is also made worse by the fact that in many radio systems, when the speech frame is defective, the decoding of the preceding speech frame is repeated by a speech decoder, so that if the preceding speech frame was defective as well, two defective speech frames are decoded successively.

[0008] Even if the channel coding could detect the errors, another problem, described in the following, may lead to an audio shock. In radio

systems, for example in the GSM system (Global System for Mobile Communication), the encryption is usually decrypted before channel decoding. However, for instance in TETRA the decryption is not performed until after the channel decoding. Hence, the situation may be such that the speech frame containing speech is actually free of defects on the basis of the channel code, too, but for some reason, for example due to incorrect input parameters, the decryption puts the contents of the speech frame that is faultless as such in complete disorder. Although the speech decoder identifies bad speech frames on the basis of the channel code, for example by receiving a bad frame indication from the channel decoder when needed, it will not be able to identify the speech frame that the decryption has put in disorder, because this kind of speech frame does not receive a bad frame indication. Hereby, the speech decoder attempts to decode the speech frame that was put in disorder, causing again, in the worst case, an audio shock. The problem is more likely to arise because of the two-level encryption option in TETRA: radio interface encryption and point-to-point encryption. Users may set up their own encryption modules to perform point-to-point decryption in TETRA mobile stations, and thus the manufacturer of the device has no opportunity to check and test the operation of this kind of encryption module, because the encryption module processes information that is already channel-decoded. From the point of view of the device manufacturer, the encryption module is a black box, a decrypted speech frame coming out of the box in an ideal case. Further, there may be errors in updating encryption keys, if an incorrect encryption key is used as an input parameter of an encryption algorithm.

25 BRIEF DESCRIPTION

[0009] An object of the invention is to provide an improved method, an improved radio system, an improved mobile station for a radio system and an improved network for a radio system. The method according to claim 1 represents one aspect of the invention. The radio system according to claim 12 represents another aspect of the invention. Still another aspect of the invention is represented by the mobile station of a radio system according to claim 23. Yet another aspect of the invention is represented by a radio system network according to claim 24. Other preferred embodiments of the invention are disclosed in the dependent claims.

[0010] The invention is based on the idea that even if the speech frame is free of defects on the basis of channel decoding, at least one speech parameter in the channel-decoded speech frame is examined, and on the basis of this examination it is decided whether it is advantageous to attempt to decode the speech frame by means of a speech decoder. Hence, an audio shock is avoided in such a way that the capability of the channel decoding to detect the errors is not blindly relied upon. On the other hand, if the decryption is performed in the system after the channel decoding, studying the parameters in the channel-decoded and decrypted speech frame also prevents a possible audio shock in a situation where the channel-decoded speech frame was free of defects but the decryption put the contents of the speech frame in disorder.

[0011] The most significant advantage of the invention is that its utilization allows avoidance of audio shocks in problematic cases described above.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Preferred embodiments of the invention will now be described by way of example with reference to the attached drawings, of which

[0013] Figure 1 shows an example of a structure of a radio system;

[0014] Figure 2 shows a structure of a radio system network and a radio system mobile station;

[0015] Figure 3 shows a structure of a radio transmitter and a radio receiver;

[0016] Figure 4 shows channel coding of speech bits;

[0017] Figure 5 shows a flow chart illustrating a method of processing a speech frame in a radio system.

DESCRIPTION OF EMBODIMENTS

[0018] The structure of a radio system is described with reference to Figure 1. The system to be described is TETRA, but it is obvious that the embodiments are not limited to this system but may be used in all radio systems that are affected by at least one of the problems mentioned above, i.e. the channel decoding being unideal and/or the decryption functioning inappropriately. The TETRA standard describes six system components and the connections between them. Internal connections of the system components are not defined in the standard in order to allow manufacturers to

implement the internal structure of each system component in the most efficient way possible. The system components are as follows: a network 100, a line station 108, a mobile station 112, a direct mode mobile station 114, 116, a gateway 104, and a network management unit 106.

5 **[0019]** The network 100 is a TETRA network system comprising separate network elements, such as a base station 212 with an antenna 210 and a mobile services switching centre 214 shown in Figure 2. The network 100 roughly corresponds to the GSM base station subsystem and network subsystem combined. This combination is indicated by reference numeral 102
10 in Figure 1. The network 100 comprises not only reference numeral 102 but also the network management unit 106 and the gateway 104. As can be seen from Figure 1, the TETRA network 100 may be connected to a second TETRA network 120, the second TETRA network 120 having a radio connection 122 to a TETRA mobile station 124.

15 **[0020]** The line station 108 is in practice a terminal or a dispatcher unit in a control room monitoring the TETRA network 100.

[0021] The gateway 104 enables calls between users of the TETRA network 100 and users of non-TETRA networks. The definition of the gateway 104 is necessary because other networks to be connected to the TETRA
20 network 100 use incompatible information formats and communication protocols, whereby translations or conversions have to be performed. As an example of a non-TETRA network, Figure 1 illustrates a public switched telephone network 130 that comprises an ordinary telephone 132, to which a connection can be established from the mobile station 112. Other non-TETRA
25 networks to which a connection can be established via the gateway 104 include ISDN (Integrated Services Digital Network) and a public data network.

[0022] By means of the network management unit 106 network management is performed locally and by telecontrol. Network management comprises for example monitoring of errors, system configuration, billing,
30 performance measurement, and planning.

[0023] The mobile station 112 enables a radio connection 110 to the network 100. The mobile station 112 can be portable or positioned in a vehicle. The direct mode mobile station 114 enables a radio connection 118 directly to another direct mode mobile station 116 without the network 100
35 being utilized in any way in implementation of the radio connection 118. The radio connection 118 may be a point-to-point connection or a multipoint

connection. An apparatus combining the ordinary mobile station 112 and the direct mode mobile station 114 is also feasible. The ordinary radio connection 110 utilizing the network 100 uses a trunking-mode air interface, whereas the radio connection 118 between two direct mode mobile stations 114, 116 uses a direct-mode air interface.

[0024] The connection between two direct mode mobile stations 140, 148 may also be implemented by using an independent radio repeater 144. Thus, a radio connection 142 between the first mobile station 140 and the radio repeater 144 uses a converted direct-mode air interface, as does a radio connection 146 between the second mobile station 148 and the radio repeater 144. Another way to extend the audibility range is to use a gateway radio repeater 152. Thus, a direct mode mobile station 156 uses a converted direct-mode air interface to implement a radio connection 154 with the gateway radio repeater 152. Radio connection 150 of the gateway radio repeater 152 is implemented with the network 100 by using a trunking-mode air interface.

[0025] Multiple encryption with different encryption keys and algorithms may be simultaneously used in one call. For example in the connection 110 between the mobile station 112 and the network 100 in Figure 1, there is radio interface encryption. There is also radio interface encryption in the connection 150 between the network 100 and the gateway radio repeater 152. Further, there is radio interface encryption in the connection 154 between the gateway radio repeater 152 and the direct-mode mobile station 156. Moreover, point-to-point encryption may be used between the mobile station 112 and the direct-mode mobile station 156. In other words, quadruple encryption may be used to implement one call. The situation could be even worse if, for example, there were a connection from the gateway radio repeater 152 to the radio repeater 144 and therefrom a further connection to the direct-mode mobile station 148. In such a case, one call could have fivefold encryption implemented with different encryption masks.

[0026] According to Figure 2, the mobile station 112 comprises an antenna 200, a transmitter-receiver 202 and a control part 204. Other components of the mobile station 112, such as means for implementing the user interface, are not described herein.

[0027] TETRA is not presented in more detail herein but reference is made to the work *Digital Mobile Communications and the TETRA System* mentioned at the beginning.

[0028] In the following, reference is made to Figure 3, which illustrates the structure of a radio receiver and a radio transmitter.

[0029] A simplified structure of a transmitter is shown at the top of Figure 3. It is obvious that the transmitter also comprises other functions and structural parts but these are not relevant in this context. Speech 320 is coded by means of a speech coder, such as the ACELP coder mentioned at the beginning, which is described in the above-mentioned work *Digital Mobile Communications and the TETRA System*. The structure of the ACELP codec is described in the ETSI (European Telecommunication Standards Institute) specification ETS 300 395-1, May 1997, *Terrestrial Trunked Radio (TETRA); Speech CODEC for full-rate traffic channel; Part 1: General Description of Speech Functions*, incorporated as reference herein.

[0030] Next, the bit stream mapping speech in the form of different parameters, produced by the speech codec, is applied to an encryption-performing block 302, in which the bit stream is encrypted in the way described at the beginning by connecting an encryption mask thereto, which encryption mask has been produced by the encryption algorithm.

[0031] The encrypted bit stream is then channel-coded in a channel coder 304; in TETRA, for example, block coding is first performed in the form of CRC, and then convolutional coding. This is shown in greater detail in Figure 4, which illustrates channel coding of speech bits. The ACELP codec produces 137 bits per each speech part of 30 milliseconds, which corresponds to a bit rate of 4,567 kbit/s. There are two times 137 bits, as two speech frames produced by the ACELP codec are positioned in one burst of the radio interface. The bits of one speech frame are divided into three different categories on the basis of the sensitivity of the bits. The 30 bits that are most sensitive to errors are placed in category 2; 56 bits are placed in category 1; and the 51 bits least sensitive to an error are placed in category 0. A cyclic redundancy check of eight bits and, in addition, four tail bits, are added to the bits of category 2. Thus, category 2 contains 72 bits that are subjected to 8/18 convolutional coding, in other words 18 coded data bits in the channel correspond to each eight data bits, in other words the final number of bits in category 2 is 162 bits. The bits in category 1 are subjected to 2/3 convolutional coding, in other words 112 data bits result in 168 coded data bits in the channel. The bits in category 0 are not protected by channel coding at all, in other words there will be the total of 102 bits of this category in the channel.

The total number of bits is thus $162 + 168 + 102 = 432$ bits, which is divided in an ordinary burst of the uplink or downlink into two fields of 216 bits.

- 5 **[0032]** In the following, the 137 bits produced by the ACELP codec are shown in Table 1. The bits form parameters that are categorized under either a filter or one of the four subframes. The filter parameters are codebook indexes, there being ten of them, LSP1 to LSP10 (LSP = Line Spectrum Pair). Each subframe contains the following parameters: pitch delay, codebook indexes for four different pulses, pulse global sign, pulse shift and a codebook index for gains.

10

Parameter category	Name of parameter	Number of bits
Filter	[0033] Codebook index: LSP1 to LSP3	8
	[0034] Codebook index: LSP4 to LSP6	9
	Codebook index: LSP7 to LSP10	9
Subframe 1	[0035] Pitch delay	8
	[0036] Codebook index: pulse 4	3
	[0037] Codebook index: pulse 3	3
	[0038] Codebook index: pulse 2	3
	[0039] Codebook index: pulse 1	5
	[0040] Pulse global sign	1
	[0041] Pulse shift	1
	Codebook index: gains	6
Subframe 2	[0042] Pitch delay	5
	[0043] Codebook index: pulse 4	3
	[0044] Codebook index: pulse 3	3
	[0045] Codebook index: pulse 2	3
	[0046] Codebook index: pulse 1	5
	[0047] Pulse global sign	1
	[0048] Pulse shift	1
	Codebook index: gains	6
Subframe 3	[0049] Pitch delay	5
	[0050] Codebook index: pulse 4	3
	[0051] Codebook index: pulse 3	3
	[0052] Codebook index: pulse 2	3
	[0053] Codebook index: pulse 1	5
	[0054] Pulse global sign	1
	[0055] Pulse shift	1
	Codebook index: gains	6
Subframe 4	[0056] Pitch delay	5
	[0057] Codebook index: pulse 4	3
	[0058] Codebook index: pulse 3	3
	[0059] Codebook index: pulse 2	3
	[0060] Codebook index: pulse 1	5
	[0061] Pulse global sign	1

	[0062] Pulse shift	1
	Codebook index: gains	6

[0063] Table 1: 137 bits produced by the ACELP codec

[0064] In the decoding of speech, the parameters shown in Table 1 are decoded and speech is reconstructed by means of a synthesis filter. The most important bits, i.e. the bits in category 2, are located in the following parameters of the speech frame:

[0065] - Codebook index: LSP1 to LSP3, speech frame bits B1 to B4;

[0066] - Codebook index: LSP4 to LSP6, speech frame bits B9 to B12;

[0067] - Codebook index: LSP7 to LSP10, speech frame bits B18 to B21;

[0068] - Pitch delay for subframe 1, speech frame bits B27 to B32;

[0069] - Codebook index in each subframe 1 to 4: gains, speech frame bits B51 to B53, B73 to B80, B105 to B107 and B132 to B134.

[0070] The encrypted and channel-coded frame is finally sent to a radio path 110 by modulating it in a modulator 306. In TETRA, the multiple access method is TDMA (Time Division Multiple Access), the modulating method being $\pi/4$ - DQPSK (Differential Quaternary Phase Shift Keying).

[0071] A simplified structure of a receiver is illustrated at the bottom of Figure 3. It is obvious that the receiver also comprises other functions and structural parts, but these are not relevant in this context. In the following, reference is also made to the flow chart of Figure 5, which illustrates a method of processing a speech frame in a radio system.

[0072] The burst received from the radio path 110 is demodulated in a demodulator 308. Channel decoding is then performed in a channel decoder 310, in other words a cyclic redundancy check is computed and the convolutional coding is decoded using a Viterbi decoder, for example. At this point, performance of a method of processing a speech frame in a radio system is actually started. Performance of the method is initiated in a block 500, and in a block 502 the channel-coded speech frame having propagated over the radio path is channel-decoded.

[0073] Subsequently, there is an optional decryption block 504 in the method, in which block the encryption of the speech frame, if used, is decrypted.

5 [0074] After this, it is checked in a block 506 whether the speech frame is free of defects on the basis of the channel-decoding. This may be performed, for example, by comparing the computed cyclic redundancy check with the eight-bit cyclic redundancy check received from the channel and contained in the bits of category 2.

10 [0075] If, according to the channel decoding, for example a cyclic redundancy check, the speech frame is defective, the block 506 is followed by a block 514, as indicated by arrow 520. As shown by the block 514, there are no attempts to correct the defective speech frame by means of a speech decoder 316 but:

[0076] A) The speech frame may be replaced with a previously received speech frame containing speech free of defects, attenuating thus the gain.

15 [0077] B) A break is put at the point of the speech frame concerned, in other words there is no speech from the loudspeaker of the apparatus for 30 milliseconds.

[0078] C) The radio connection 110 can be released if it is too bad, for example if there have been defective frames during a certain period of time, or a certain number of defective frames successively.

20 [0079] If the call terminates, the block 514 is followed by a block 516, as indicated by arrow 530, the performance of the method being terminated in this block. If the call continues, the block 514 is followed by the block 502, as indicated by arrow 528 to process the next received speech frame.

25 [0080] If the speech frame is free of defects on the basis of the channel decoding, the block 506 is followed by a block 508, as indicated by arrow 522. In the block 508, it is inferred from the value of at least one speech parameter in the channel-decoded speech frame whether the speech frame contains speech that is decodable by means of a speech decoder. In TETRA,
30 for example, a speech parameter refers to parameters produced by an ACELP codec, as was shown in Table 1. This inference can be performed as described in US patent 6,021,385 (Järvinen et al.) by forming a probability distribution for the values of a speech parameter, with the difference that the method is not taken into use based on the quality of the radio link but on the
35 success of the channel decoding. A difference is also that speech frame bits protected by channel coding and most sensitive to errors are used in the

inference 508, in other words TETRA also utilizes bits of category 2 in the inference of the block 508. Still another difference in the method of the present application compared with said US patent is that the US patent cannot solve the problems described in this application. The object of US 6,021,385 is to
 5 improve the quality of speech decoding, and it utilizes information about the quality of a radio link, for example reception power or bit error ratio.

[0081] The result of the inference of the block 508 is checked in a block 510. If, on the basis of the inference, the speech frame does contain speech that is decodable by means of the speech decoder 316, the following
 10 step is block 512, in which the speech frame is decoded by the speech decoder 316. If the call terminates, the block 512 is followed by a block 516, as indicated by arrow 532, the implementation of the method being terminated in this block. If the call continues, the block 512 is followed by the block 502, as indicated by arrow 534, to process the next received speech frame. If,
 15 according to the inference, the speech frame does not contain speech that would be decodable by means of the speech decoder 316, the next step is the above-described block 514, as indicated by arrow 524, whereby the speech frame is not decoded but one of the replacement measures of block 514 is taken.

[0082] The radio system according to Figure 3 thus comprises also inferring means 314 for inferring from the value of at least one speech parameter in the channel-decoded speech frame whether the speech frame contains speech that is decodable by the speech decoder 316 if the speech frame is free of defects according to the channel decoder 310. Further, the
 20 speech decoder 316 is arranged to decode the speech frame if, according to the inference, the speech frame does contain speech that is decodable by the speech decoder 316, and the speech decoder 316 is arranged not to decode the speech frame if, according to the inference, the speech frame does not contain speech that would be decodable by the speech decoder. A potential
 25 decryption device 312 is connected between the channel decoder 310 and the inferring means 314, even if in systems other than TETRA the decryption device may be located before the channel decoder 310. From the decryption device 312 the decrypted speech frame is taken to the inferring means 314 and, as indicated by arrow 324, also to the speech decoder 316.

[0083] As indicated by arrow 322, the channel decoder 310 sends a bad frame indication to the speech decoder 316 if, on the basis of the channel

decoding, the speech frame is defective. In an optimal case, speech 320 that sounds the same as the speech originally sent from the transmitter comes out of the speech decoder 316.

[0084] In a preferred embodiment of the invention, the inferring means 314 are arranged to send a bad frame indication 324 to the speech decoder 316 if the speech frame does not, according to the inference, contain speech that would be decodable by means of the speech decoder 316. Another solution that can be used to inform the speech decoder 316 that it is not advantageous to attempt to decode the speech frame is such that the inference means 314 are arranged to send a homing sequence 326 to the speech decoder 316 if, according to the inference, the speech frame does not contain speech that would be decodable by the speech decoder 316. The purpose of the homing sequence is to bring the speech decoder 316 into a known state and thus to enable such an output for the loudspeaker which will not cause an audio shock.

[0085] In a preferred embodiment, the inference according to the method in the block 508 is performed in such a way that the inferring means 314 perform the inference by using probability calculation. This is implemented in such a way, for example, that the inferring means 314 calculate in the inference the probability of the value of at least one speech parameter. If the value of a speech parameter is highly improbable in such a situation, it can be assumed that the speech frame does not contain speech. Naturally, the use of two or more parameters for which probability is calculated increases the reliability of the inference. A value may also be calculated for the probability of change in the value of a speech parameter. In such a case, a threshold value has been defined in the inferring means 314 for the probability of change in the value of the parameter to be compared during a given number of speech frames. If the probability of change is lower than the threshold, the inferring means 314 infer that the speech frame does not contain speech that would be decodable by the speech decoder 316. To take an example, the pitch of the speaker does not normally change much during a call. If the value of the pitch-mapping parameter changes too much between two successive speech frames, it can be inferred that the speech frame is corrupted either because of disturbance on the radio path and/or erroneous functioning of decryption. Thus, the interest lies not in the parameter value or in the change therein as

such but in the probability of the parameter value or the probability of change in the parameter value.

[0086] In principle, the point of the probability calculation for parameters is that in reality there is always correlation between the parameter bits. Probability calculation may be used to calculate probabilities between
 5 parameters within one speech frame or between some parameters in successive speech frames. Both calculation methods may naturally be used simultaneously to confirm the result, for example by first calculating probabilities between speech frames and then, if a speech frame begins to
 10 look corrupted, a probability is calculated for at least one parameter in the speech frame concerned. Comparing successive speech frames requires memory in connection with the inferring means 314, in which memory preceding calculated values or whole speech frames can be stored.

[0087] The transmitter/receiver pair according to Figure 3 may be
 15 composed of for example two ordinary TETRA mobile stations 112; two TETRA direct mode mobile stations 114; a mobile station 112 and a network 100; or a network 100 and a mobile station 112. When the connection is between the network 100 and a mobile station, speech may be decoded in the gateway 104 prior to entering a non-TETRA network 130, or in some part of
 20 the network 100, such as in a mobile services switching centre 214 prior to entering another TETRA network 120 or a line station 108.

[0088] The required functionality may be implemented for example as software in a general-purpose processor, whereby the required functions are implemented as software components. A hardware implementation may
 25 also be used, for example as ASIC (Application Specific Integrated Circuit) or as control logic constructed of separate components. Typically, all functions shown in Figure 3 may be implemented as software. Thus, the inferring means 314 significant for the invention may also be advantageously implemented as software which receives input data from the decryption device 312 and
 30 comprises an input connection 326 to the speech decoder 316. Software may be performed for example in the control part 204 of the mobile station 112, and in the network 100 for example in the control part of the mobile services switching centre 214 or the gateway 104.

[0089] While the invention is described above with reference to the
 35 example according to the attached drawings, it is obvious that the invention is

not confined thereto but may be modified in many ways within the inventive idea disclosed in the attached claims.

104760 2024550